

Dynamic Resource Allocation in Next Generation Cellular Networks with Full-Duplex Self-backhauls

Lei Chen*, F. Richard Yu*, Hong Ji[†], Bo Rong[‡], and Victor C.M. Leung[§]

*Depart. of Systems and Computer Eng., Carleton Univ., Ottawa, ON, Canada

[†]Key Lab. of Universal Wireless Comm., Beijing Univ. of Posts and Telecom., P.R. China

[‡]Communications Research Centre, Ottawa, ON, Canada

[§]Depart. of Electrical and Computer Eng., The Univ. of British Columbia, Vancouver, BC, Canada

Abstract—With the dense deployment of small cell networks, low-cost backhaul schemes for small cell base stations (SBSs) have attracted great attentions. Self-backhaul using cellular communication technology is considered as a promising solution. Although some excellent works have been done on self-backhaul in small cell networks, most of them do not consider the recent advances of full-duplex (FD) and massive multiple-input and multiple-output (MIMO) technologies. In this paper, we propose a self-backhaul scheme for small cell networks by combining FD and massive MIMO technologies. In our proposed scheme, the macro base station (MBS) is equipped with massive MIMO antennas, and the SBSs have the FD communication ability. By treating the SBSs as *special* macro users, we can achieve the simultaneous transmissions of the access link of users and the backhaul link of SBSs in the same frequency. Furthermore, considering the existence of inter-tier and intra-tier interference, we formulate the power allocation problem of the MBS and SBSs as an optimization problem. Because the formulated power allocation problem is a non-convex problem, we transform the original problem into a difference of convex program (DCP) by successive convex approximation method (SCAM) and variable transformation, and then solve it using a constrained concave convex procedure (CCCP) based iterative algorithm. Finally, extensive simulations are conducted with different system configurations to verify the effectiveness of the proposed scheme.

Index Terms—Small cell networks, self-backhaul, full duplex, massive MIMO

I. INTRODUCTION

With the explosive demand for mobile broadband services and the emergence of new high capacity mobile devices, mobile networks have to continuously evolve to meet capacity and coverage demands with the latest technologies. At the same time, there is an apparent trend of declining profitability of mobile data despite the recent exponential growth of mobile data usage [1]. Small cell networks, which can improve spectrum efficiency, energy efficiency, and coverage effectively, as well as reduce the capital expenses (CapEx) and operation expenses (OpEx), have been considered as an important technology of next generation cellular networks [2]–[4]. Furthermore, dense small cell networks have attracted great attentions, where a mass of small cell base stations (SBSs) are deployed to improve the quality of service (QoS) further [5]. However,

with the dense deployment of massive SBSs, the *backhaul* problem becomes severe increasingly. Therefore, optimizing backhaul is critical to satisfy the QoS of users and reduce the CapEx and OpEx.

The backhaul technologies for small cells can be classified into three categories. The first is the wired optical fibre, which has high capacity and will undoubtedly connect a major portion of small cells, especially in the long run [6]. The second is wireless point-to-point microwave [7] or mmWave [8], which uses high-gain directional antennas in line-of-sight (LOS) environments and provides high-capacity backhaul link. Unfortunately, installing fibres, microwave or mmWave equipment is expensive and time-consuming, preventing fast deployment of SBSs [9]. In addition, the microwave and mmWave usually operate in the LoS environment and are not suitable for the urban environment because of the massive buildings. The third possible backhaul technology is the cellular communication technology (e.g., LTE). It uses the cellular spectrum to access and backhaul, and is suitable for the non-LoS environment because of the radio nature of cellular spectrum [10]. Moreover, with the cellular communication technology backhaul, which is called *self-backhaul* in the literature (e.g., [11]), the SBSs do not require extra backhaul equipment or spectrum, and consequently, self-backhaul is a promising technology in future small cell networks [11].

Nevertheless, due to the limited cellular spectrum and the existence of inter-tier and intra-tier interference, designing self-backhaul schemes is challenging. In [12], the authors proposed a multi-hop self-backhaul scheme by jointly considering resource allocation and routing. The authors of [13] studied the fair scheduling problem in wireless multi-hop self-backhaul networks. The synchronization issue of time division duplex (TDD)-based self-backhaul was studied in [10]. Although those works are excellent, they focus on the traditional half-duplex self-backhaul. With the development of self-interference cancellation technologies, full-duplex (FD) communication technology becomes possible. FD makes it possible for radios to transmit and receive simultaneously in the same frequency band, which nearly doubles the spectrum efficiency [14]. The FD relay has been researched in many papers [15] [16] and the authors of [17] first proposed FD self-backhaul scheme where FD communication hardware is equipped in SBSs. Consequently, SBSs receive data from the macro base station (MBS) and transmit it to SBSs in downlink

This paper is jointly supported by the Hi-Tech Research and Development Program of China (National 863 Program) under Grant 2014AA01A701 and National Natural Science Foundation of China under Grant 61271182.

(DL); at the same time, they receive data from mobile users in uplink (UL) and transmit it to the MBS and in same frequency. Moreover, the effectiveness of FD self-backhaul scheme had been proved in this paper. However, the authors do not consider the interference problem in small cell networks. In particular, massive MIMO technology is not involved in this paper.

Massive MIMO technology can achieve the transmission of multiple users at the same time and in the same frequency band by utilizing the large spatial degrees-of-freedom, which will also improve the spectrum efficiency of the system [18]. Due to the fact that the MBS is responsible for the transmission to SBSs besides its users in FD self-backhaul scheme, the MBS with massive MIMO can transmit or receive data to SBSs and its users simultaneously in the same frequency band. In this way, the spectrum efficiency of backhaul and access links will be improved significantly. In [19], [20], the authors intended to analyze the gain of jointly considering massive MIMO and small cells, but they did not consider the backhaul problem of small cell networks and simplified the power allocation of MBS and SBSs by equal power allocation.

Despite the potential vision of FD self-backhaul small cell networks with massive MIMO, many research challenges remain to be addressed. One of the main research challenges is resource allocation, which plays an important role in traditional wireless networks [21]–[30]. When FD self-backhaul and massive MIMO are jointly considered, the problem of resource allocation becomes even more challenging. On one hand, the FD self-backhaul makes the backhaul and access links coupled, which depends on power allocation and self-interference cancellation performance. On the other hand, massive MIMO will introduce the interference among multiple users. To the best of our knowledge, the problem of power allocation in FD self-backhaul small cell networks with massive MIMO has not been studied in previous works. The distinct features of this paper are summarized as follows

- We propose a novel architecture of FD self-backhaul small cell networks with massive MIMO technology. In our proposed scheme, the MBS is equipped with massive MIMO antennas, and the SBSs have the FD capability. By treating the SBSs as *special* macro users, we can achieve the transmissions of the access link of users and the backhaul link of SBSs simultaneously in the same frequency band.
- Furthermore, considering the existence of inter-tier and intra-tier interference, we formulate the power allocation problem of the MBS and SBSs as an optimization problem, which maximizes the total spectrum efficiency (SE) of the small cell networks, while considering the lowest quality of service (QoS) requirement of users. In addition, we take the residual self-interference of FD communications into account in the formulated problem.
- Since the formulated problem is a non-convex optimization problem, its computational complexity is high. To solve it efficiently, we transform the original problem into a difference of convex program (DCP) by using successive convex approximation method (SCAM) and appropriate variable substitution, and then solve it using

a constrained concave convex procedure (CCCP)-based iterative algorithm, which reduces the computation complexity significantly. Furthermore, we prove the convergence of our proposed iteration algorithm.

- Extensive simulations are conducted with different system configurations to verify the effectiveness of the proposed FD self-backhaul scheme with massive MIMO.

The rest of this paper is organized as follows. The proposed FD self-backhaul scheme with massive MIMO is described in Section II. The power allocation problem is formulated in Section III. Then we solve the optimization problem in Section IV. Simulation results are discussed in Section V. Finally, we conclude this study and describe our future works in Section VI.

II. NETWORK MODEL

In this paper, we consider a two-tier small cell network consisting of one MBS with M antennas and N SBSs with single antenna¹, as shown in Fig. 1. To fully exploit the spectrum resource, the MBS and SBSs share the spectrum. For ease of presentation, we assume that the MBS serves K single-antenna users (MUs) and each SBS serves one single-antenna user (SU). Note that $K+N \ll M$, which means that a massive MIMO system is adopted in this paper and M could be very large (e.g., 100, 1000, or even more [31]). We define by the set \mathcal{U}^m and the set \mathcal{U}^s the MUs and SUs, respectively, where the element u_k^m of \mathcal{U}^m and u_n^s of \mathcal{U}^s represent the k -th MU and the SU of the n -th small cell BS, respectively. To reap the benefits of massive MIMO antennas, channel state information (CSI) must be available at transmitter, the TDD protocol is adopted in this paper because the channel reciprocity can be exploited², which allows the MBS to estimate its DL channel from UL pilots sent by the users. In this paper, we assume that all base stations (BSs) serve users over flat-fading channel.

A. FD Self-backhaul Scheme of Small Cell Networks with Massive MIMO

As shown in Fig. 2(a), SBSs are equipped with FD hardware, which enables them to backhaul data for themselves. In the DL, a SBS can receive data from the MBS while simultaneously transmitting to its users at the same resource block. In the UL, a SBS can receive data from the users while simultaneously transmitting data to the MBS at the same resource block. Note that the In this mechanism, the small cell can effectively backhaul itself, eliminating the need for a separate backhaul solution or a separate backhaul frequency band. Therefore, self-backhauling can significantly reduce the cost and complexity of rolling out small cell networks. In order to distinguish DL from UL in access and backhaul transmissions, we call the relevant links as access UL, access DL, backhaul UL, and backhaul DL, respectively. Due to the limitation of self-interference cancellation technologies, the

¹To simplify the network model, single-antenna SBS is assumed in this paper and the works of this paper can be expanded to small cell networks with multi-antenna SBS by involving multi-antenna channel model.

²The proposed FD self-backhaul scheme with massive MIMO also can be used in FDD system based on some channel estimation method [32], [33].

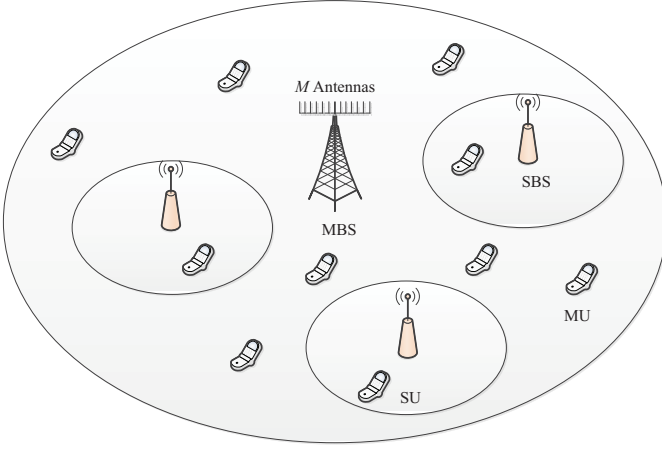


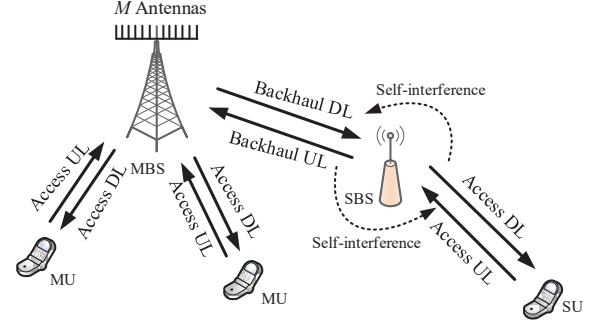
Fig. 1: Network framework.

backhaul DL and access UL will suffer some self-interference from access DL and backhaul UL, respectively.

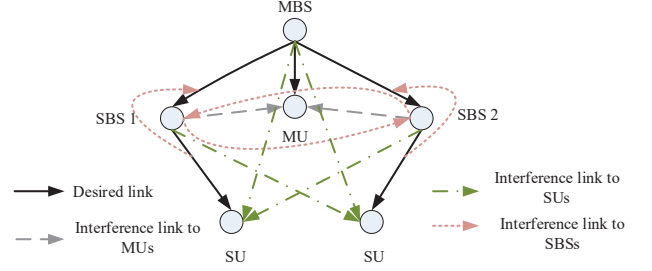
In massive MIMO systems, the BS can transmit or receive data to/from multiple users simultaneously in the same frequency band by using the beamforming technology or by using virtual MIMO technology, respectively. We study the FD self-backhaul in small cell networks with massive MIMO, where the SBSs are considered as *special MUs*. In the DL, the MBS transmits data to MUs and SBSs, and SBSs transfer the data to their users simultaneously in the same frequency band. In the UL, the SBSs receive data from their users and transfer them to the MBS, and the MBS receives data from MUs and SBSs simultaneously in the same frequency band. In other words, by jointly considering FD and massive MIMO technologies, we achieve not only the transmission or reception of MUs and SUs in the same frequency band at the same time, but also the access and backhaul of SBSs in the same frequency band at the same time. This scheme will improve the spectrum efficiency and decrease the cost of backhaul infrastructure. Comparing with DL, UL usually has less traffic, so we focus on the transmission of DL in this paper. For ease of analysis, full buffer traffic model is assumed in the MBS and SBSs. Note that pilot contamination problem in massive MIMO system is not considered due to the fact that many excellent works have been done.

B. Channel Model

The channel matrix from the transmission node to reception node can be written as $\mathbf{G} = \mathbf{D}^{1/2}\mathbf{H}$, where $\mathbf{D} = \text{diag}\{\beta_1, \beta_2, \dots, \beta_I\}$ (I indicates the number of reception nodes, $I = 1$ for SBSs' DL and $I = K + N$ for MBS' DL). The component $\beta_i = \varphi\zeta/d_i^\alpha$ consists of path loss and shadow fading, φ is a constant related to carrier frequency and antenna gain, d_i is the distance between the two nodes, α is the path loss exponent, ζ represents the shadow fading which follows the log-normal distribution $10 \log \zeta \sim N(0, \sigma^2)$. Fast fading matrix is $\mathbf{H} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_I^T]^T \in C^{I \times J}$ (J represents the number of antennas of transmission node, $J = 1$ for SBSs' DL and $J = M$ for MBS' DL), the components $h \sim N(0, 1)$



a) The FD self-backhaul scheme with massive MIMO



b) The DL interference graph

Fig. 2: The FD self-backhaul scheme with massive MIMO and the DL interference graph.

TABLE I: Notation

Notation	Definition
K	The number of MUs
M	The number of the MBS' antennas
N	The number of SBSs
P	Transmission power
β	Large scale fading
\mathbf{h}	Small scale fading
\mathbf{w}	Precoding matrix
s	Transmission symbol
k	The index of the k -th MU
n	The index of the n -th SBS and its user
$(*)^m$	The related variables of MBS DL
$(*)^b$	The related variables of backhaul DL
$(*)^s$	The related variables of SBS DL

are Rayleigh flat-fading random variables. In this paper, the Zero-Forcing Beamforming is adopted for MBS DL. With it, the multi-user interference can be eliminated perfectly [34]. To reduce the complexity and make beamforming scheme effective, the Zero-Forcing Beamforming only be operated among MUs' DLs without including the backhaul DLs. $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{K+N}] \in C^{M \times (K+N)}$ is defined as the precoding matrix of the MBS. For easy reading, the notations of this paper are described as in Table I.

For MUs, they receive signal from the MBS while suffering the inter-tier interference from SBSs, so the received signal of

MU k is given as

$$y_k^m(t) = \underbrace{\sqrt{P_k^m \beta_k^m} \mathbf{h}_k^m \mathbf{w}_k^m s_k^m}_{\text{desired signal}} + \underbrace{\sum_{n=1}^N \sqrt{P_n^s \beta_{nk}^s} h_{nk}^s s_n^s}_{\text{inter-tier interference}} + w_k^m. \quad (1)$$

Based on the equation above, the received signal to interference and noise ratio (SINR) of MU k can be written as

$$\xi_k^m = \frac{P_k^m \beta_k^m \|\mathbf{h}_k^m \mathbf{w}_k^m\|^2}{\sum_{n=1}^N P_n^s \beta_{nk}^s h_{nk}^s{}^2 + \sigma^2}. \quad (2)$$

For SBSs, they receive data from the MBS and forward the data to SUs simultaneously in the same frequency band, which results in that the SBSs suffer not only the intra-tier interference from other SBSs but also the *self-interference* from themselves. As a result, the received signal of SBS n can be expressed as

$$y_n^b(t) = \underbrace{\sqrt{P_n^b \beta_n^b} \mathbf{h}_n^b \mathbf{w}_n^b s_n^b}_{\text{desired signal}} + \underbrace{\sum_{n'=1, \neq n}^N \sqrt{P_{n'}^b \beta_{n'n}^b} h_{n'n}^b s_{n'}^b}_{\text{intra-tier interference}} + \underbrace{\sqrt{\gamma P_n^s s_n^s}}_{\text{self-interference}} + w_n^b, \quad (3)$$

where $\sqrt{\gamma P_n^s s_n^s}$ represents the self-interference signal and the value of γ is determined by self-interference cancellation technologies. Any self-interference cancellation technology (e.g., [35] and [11]) can be applied at the SBSs, and the analysis in this paper is a general case. Thus, the received SINR of SBS n can be written as

$$\xi_n^b = \frac{P_n^b \beta_n^b \|\mathbf{h}_n^b \mathbf{w}_n^b\|^2}{\sum_{n'=1, \neq n}^N P_{n'}^b \beta_{n'n}^b h_{n'n}^b{}^2 + \gamma P_n^s + \sigma^2}. \quad (4)$$

For SUs, they receive signal from their associated SBSs while suffering the inter-tier interference from the MBS and the intra-tier interference from other SBSs. Mathematically, the received signal of n -th SBS' user can be expressed as

$$y_n^s(t) = \underbrace{\sqrt{P_n^s \beta_n^s} h_n^s s_n^s}_{\text{desired signal}} + \underbrace{\sum_{n'=1, \neq n}^N \sqrt{P_{n'}^s \beta_{n'n}^s} h_{n'n}^s s_{n'}^s}_{\text{intra-tier interference}} + \underbrace{\sum_{k=1}^K \sqrt{P_k^m \beta_{nk}^m} \mathbf{h}_n^m \mathbf{w}_k^m s_k^m + \sum_{n'=1}^N \sqrt{P_{n'}^b \beta_{n'n}^b} \mathbf{h}_n^m \mathbf{w}_{n'}^b s_{n'}^b}_{\text{inter-tier interference}} + w_n^s. \quad (5)$$

Then, its received SINR can be written as

$$\xi_n^s = \frac{P_n^s \beta_n^s h_n^s{}^2}{I_{\text{inter}} + \sum_{n'=1, \neq n}^N P_{n'}^s \beta_{n'n}^s h_{n'n}^s{}^2 + \sigma^2}, \quad (6)$$

where

$$I_{\text{inter}} = \sum_{k=1}^K P_k^m \beta_{nk}^m \|\mathbf{h}_n^m \mathbf{w}_k^m\|^2 + \sum_{n'=1}^N P_{n'}^b \beta_{n'n}^b \|\mathbf{h}_n^m \mathbf{w}_{n'}^b\|^2. \quad (7)$$

Based on the Shannon capacity

$$R = \log\{1 + \text{SINR}/\omega\}, \quad (8)$$

the rate of access DL of MUs, the backhaul DL of SBSs and the access DL of SUs are defined by R_k^m , R_n^b , and R_n^s , respectively. $\omega = 2 \ln(5P_e)/3$ is the SINR gap between Shannon channel capacity and a practical modulation and coding scheme achieving the BER P_e [36]. The total rate of all users can be expressed as

$$\mathcal{R}(\mathbf{P}) = \sum_{k=1}^K R_k^m + \sum_{n=1}^N R_n^s, \quad (9)$$

where $\mathbf{P} = [P_1^m, \dots, P_K^m, P_1^b, \dots, P_N^b, P_1^s, \dots, P_N^s]$ is the power allocation scheme of the MBS and SBSs.

III. PROBLEM FORMULATION

In small cell networks, power allocation is an important research issue because a power allocation scheme decides the network interference (including inter-tier and intra-tier interference) level. Different from traditional wired backhaul schemes, in our proposed FD backhaul scheme, extra power is needed in the MBS to support the backhaul link. The power allocated to SBSs' backhaul link not only affects the spectrum efficiency of MUs but also decides the spectrum efficiency of SBSs. As a result, power allocation is critical in our proposed FD self-backhaul scheme of small cell networks.

The optimal power allocation policy \mathbf{P}^* of the MBS and SBSs can be obtained by solving

$$\max_{\mathbf{P}} \mathcal{R}(\mathbf{P}) \quad (10)$$

$$\begin{aligned} \text{s.t. } & C1: R_n^s \leq R_n^b, \forall n, \\ & C2: \sum_{l=1}^{K+S} P_l^m \leq P_{\max}^m, \\ & C3: P_n^s \leq P_{\max}^s, \forall n, \\ & C4: R_k^m \geq R_{\min}, \forall k, \\ & C5: R_n^s \geq R_{\min}, \forall n. \end{aligned}$$

$C1$ specifies that the rate of SBSs' backhaul DL must be no less than that of SBS' access DL to satisfy the quality of service (QoS) of SUs. $C2$ is a transmission power constraint for the MBS in the DL and P_{\max}^m is the maximum transmission power of MBS. Similarly, $C3$ is the transmission power constraint of SBSs and P_{\max}^s is the maximum transmission power of small cell BSs. $C4$ and $C5$ specify the lowest QoS requirements of MUs and SUs, respectively.

IV. SOLUTION TO THE OPTIMIZATION PROBLEM

The problem in (10) is a non-convex problem because of the non-convexity of the objective function and the feasible set. As a result, finding a global optimum of problem (10) is computationally expensive or even intractable. In this case, designing low-complexity algorithms to compute local optimal of problem (10) is more meaningful in practice. In this section, we firstly reformulate the original problem (10) into an equivalent DCP by using SCAM and an appropriate transformation

of variables. Then a low-complexity CCCP-based algorithm is proposed to solve the DCP. To make the CCCP-based algorithm more effective, we propose an initial point searching algorithm to assist it. Finally, we conclude our proposed power allocation algorithm and prove its convergence.

A. Re-formulation of the Original Problem

According to [37], due to the fact that constant 1 exists in the Shannon Equation (8), it is easy to lead to the objective function of (10) being the difference of concave functions. In that way, it is difficult to solve the problem (10). In order to avoid the difference of concave structure in objective function of (10), we leverage the SCAM in [37] and extend this procedure to deal with (10). We make use of the following lower bound

$$\alpha \log z + \mu \leq \log(1 + z) \quad (11)$$

that is tight with equality at a chosen value z_0 when the approximation constants α, μ are chosen as.

$$\begin{aligned} \alpha &= \frac{z_0}{1 + z_0}, \\ \mu &= \log(1 + z_0) - \frac{z_0}{1 + z_0} \log z_0. \end{aligned} \quad (12)$$

Applying (11) and transformation $\tilde{p} = \ln p$ [38], [39] to (8) results in the R_k^m , R_n^b , and R_n^s relaxing to (13), (14) and (15), respectively, where all the values of α and μ are fixed, $a = 1/\ln 2$ is from the change of base of logarithms. Then we define

$$\bar{\mathcal{R}}(\tilde{\mathbf{P}}) = \sum_{k=1}^K \bar{R}_k^m(\tilde{\mathbf{P}}) + \sum_{n=1}^N \bar{R}_n^s(\tilde{\mathbf{P}}). \quad (16)$$

Furthermore, the optimization problem (10) can be relaxed as

$$\begin{aligned} \max_{\tilde{\mathbf{P}}} \quad & \bar{\mathcal{R}}(\tilde{\mathbf{P}}) \\ \text{s.t.} \quad & C1 : \bar{R}_n^b(\tilde{\mathbf{P}}) - \bar{R}_n^s(\tilde{\mathbf{P}}) \triangleq \varphi_n(\tilde{\mathbf{P}}) \geq 0, \forall n \\ & C2 : \sum_{k=1}^K e^{\tilde{\mathbf{P}}_k^m} + \sum_{n=1}^N e^{\tilde{\mathbf{P}}_n^b} \leq P_{\max}^m, \\ & C3 : e^{\tilde{\mathbf{P}}_n^s} \leq P_{\max}^s, \forall n, \\ & C4 : R_k^m(\tilde{\mathbf{P}}) \geq R_{\min}, \forall k, \\ & C5 : R_n^s(\tilde{\mathbf{P}}) \geq R_{\min}, \forall n. \end{aligned} \quad (17)$$

It is obvious that

$$\bar{\mathcal{R}}(\tilde{\mathbf{P}}) \leq \mathcal{R}(\tilde{\mathbf{P}}), \quad (18)$$

which means that problem (17) is maximizing the lower bound of the objective function of the problem (10). Therefore, it is natural to iteratively tighten the bound by updating the choice of α and μ according to the new SINR values by setting

$$[z_0]^t = [\xi(\tilde{\mathbf{P}})]/\omega]^t \quad \forall \xi_k^m, \xi_n^b, \xi_n^s, \quad (19)$$

where t is the iteration indicator. In each iteration, we just need to solve problem (17). In this way, the original problem is simplified. In order to deal with problem (17), we give the following Lemma

Lemma 1: Given α and μ , $\bar{R}_k^m(\tilde{\mathbf{P}})$, $\bar{R}_n^b(\tilde{\mathbf{P}})$ and $\bar{R}_n^s(\tilde{\mathbf{P}})$ are jointly concave with respect to the optimization variable $\tilde{\mathbf{P}}$.

Proof: Because of the similar structure among $\bar{R}_k^m(\tilde{\mathbf{P}})$, $\bar{R}_n^b(\tilde{\mathbf{P}})$ and $\bar{R}_n^s(\tilde{\mathbf{P}})$, we take $\bar{R}_n^s(\tilde{\mathbf{P}})$ as an example considering that the structure of $\bar{R}_n^s(\tilde{\mathbf{P}})$ is most complex. We define $f(\mathbf{x}) = \log \sum_{k=1}^K e^{x_k}$, which decides the convexity of $\bar{R}_n^s(\tilde{\mathbf{P}})$. Referring to [40](Sec. 3.2.6), we know that $f(\mathbf{x})$ is jointly convex in variables x_k . Due to the fact that $\bar{R}_n^s(\tilde{\mathbf{P}})$ is the sum of a linear function and some concave functions, it is jointly concave with respect to the optimization variable $\tilde{\mathbf{P}}$. The proofs of $\bar{R}_k^m(\tilde{\mathbf{P}})$ and $\bar{R}_n^b(\tilde{\mathbf{P}})$ are the same.

Based on Lemma 1, we know that the objective function of problem (17) is a strictly concave with respect to $\tilde{\mathbf{P}}$, since it is a sum of concave terms. What's more, the feasible set (defined by \mathcal{F}) decided by $C2, C3, C4$ and $C5$ is a convex set, since $C2$ and $C3$ are convex and $C4$ and $C5$ are concave. However, function $\varphi_n(\tilde{\mathbf{P}})$ represents the difference of two concave functions, i.e., a difference of convex decomposition [41] [42]. Thus, the problem (17) can be seen as a DCP, which can be rewritten as

$$\max_{\tilde{\mathbf{P}}} \quad \bar{\mathcal{R}}(\tilde{\mathbf{P}}) \quad \text{s.t.} \quad C1, \quad \tilde{\mathbf{P}} \in \mathcal{F}. \quad (20)$$

B. The Proposed Iterative Solution Algorithm

In this subsection, the CCCP, which is widely adopted for solving DCP [43] [44], is used to solve problem (20). The main idea of the CCCP based algorithm is to iteratively approximate the original nonconvex feasible set decided by $C1$ by a convex subset and then solve the resulting convex approximation in each iteration. As the nonconvex part in problem (20) stems from the fact that function $\bar{R}_n^s(\tilde{\mathbf{P}})$ is concave but not convex, we approximate this function in the t -th iteration by its first-order Taylor expansion $\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}})$ around the current point $[\tilde{\mathbf{P}}]^t$. According to [45], the first-order Taylor expansion \hat{R}_n^s is given by

$$\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}}) = \bar{R}_n^s([\tilde{\mathbf{P}}]^t) + \Delta \bar{R}_n^s([\tilde{\mathbf{P}}]^t)(\tilde{\mathbf{P}} - [\tilde{\mathbf{P}}]^t), \forall n, \quad (21)$$

which is an affine function about $\tilde{\mathbf{P}}$. Here, $\Delta \bar{R}_n^s([\tilde{\mathbf{P}}]^t)$ denotes the first-order derivative of the function $\bar{R}_n^s(\tilde{\mathbf{P}}^t)$ with respect to vector $\tilde{\mathbf{P}}$.

Then, in the t -th iteration of the proposed CCCP based iterative algorithm, the following convex optimization problem,

$$\begin{aligned} \max_{\tilde{\mathbf{P}}} \quad & \bar{\mathcal{R}}(\tilde{\mathbf{P}}) \\ \text{s.t.} \quad & C1 : \bar{R}_n^b(\tilde{\mathbf{P}}) - \hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}}) \geq 0, \forall n, \\ & \tilde{\mathbf{P}} \in \mathcal{F}, \end{aligned} \quad (22)$$

is solved, and the solution is denoted by $[\tilde{\mathbf{P}}]^{t+1}$. This procedure is carried out iteratively until convergence or until the maximum number of allowable iterations is reached. Since $\bar{R}_n^s(\tilde{\mathbf{P}})$ is concave and is approximated by its first-order Taylor expansion $\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}})$ in problem (22), it is obvious that

$$\bar{R}_n^s(\tilde{\mathbf{P}}) \leq \hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}}), \forall n, \quad (23)$$

which implies that

$$\bar{R}_n^b(\tilde{\mathbf{P}}) - \bar{R}_n^s(\tilde{\mathbf{P}}) \geq \bar{R}_n^b(\tilde{\mathbf{P}}) - \hat{R}_n^s(\tilde{\mathbf{P}}). \quad (24)$$

$$\begin{aligned}\bar{R}_k^m(\tilde{\mathbf{P}}) &= \alpha_k^m \log\{\xi_k^m(\tilde{\mathbf{P}})/\omega\} + \mu_k^m \\ &= \alpha_k^m * \left\{ a\tilde{P}_k^m + \log\{\beta_k^m \|\mathbf{h}_k^m \mathbf{w}_k^m\|^2\} - \log\left\{ \sum_{n=1}^N e^{\tilde{P}_n^s} \beta_{kn}^m h_{kn}^m{}^2 + \sigma^2 \right\} - \log\{\omega\} \right\} + \mu_k^m,\end{aligned}\quad (13)$$

$$\begin{aligned}\bar{R}_n^b(\tilde{\mathbf{P}}) &= \alpha_n^b \log\{\xi_n^b(\tilde{\mathbf{P}})/\omega\} + \mu_n^b \\ &= \alpha_n^b * \left\{ a\tilde{P}_n^b + \log\{\beta_n^b \|\mathbf{h}_n^b \mathbf{w}_n^b\|^2\} - \log\left\{ \sum_{n'=1, \neq n}^N e^{\tilde{P}_{n'}^s} \beta_{n'n}^b h_{n'n}^b{}^2 + \mu e^{\tilde{P}_n^s} + \sigma^2 \right\} - \log\{\omega\} \right\} + \mu_n^b,\end{aligned}\quad (14)$$

$$\begin{aligned}\bar{R}_n^s(\tilde{\mathbf{P}}) &= \alpha_n^s \log\{\xi_n^s(\tilde{\mathbf{P}})/\omega\} + \mu_n^s \\ &= \alpha_n^s * \left\{ a\tilde{P}_n^s + \log\{\beta_n^s h_n^s{}^2\} - \log\left\{ \sum_{k=1}^K e^{\tilde{P}_k^m} \beta_n^m \|\mathbf{h}_n^m \mathbf{w}_k^m\|^2 + \sum_{n'=1}^N e^{\tilde{P}_{n'}^b} \beta_n^m \|\mathbf{h}_n^m \mathbf{w}_{n'}^b\|^2 \right. \right. \\ &\quad \left. \left. + \sum_{n'=1, \neq n}^N e^{\tilde{P}_{n'}^s} \beta_{n'n}^s h_{n'n}^s{}^2 + \sigma^2 \right\} - \log\{\omega\} \right\} + \mu_n^s,\end{aligned}\quad (15)$$

Algorithm 1 The Proposed Low-Complexity Solution

- 1: **Initialization:** Initialize maximum number of iterations T_2^{\max} and the maximum tolerance ϵ_2 ; Initialize the algorithm with a feasible point $[\tilde{\mathbf{P}}]^0$ and set the iteration number $t=0$.
- 2: **Repeat:**
- 3: Compute the affine approximation $\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}})$ according to (21).
- 4: Solve problem (22), and update $[\tilde{\mathbf{P}}]^{t+1}$.
- 5: Set $t = t + 1$.
- 6: **Until:** $\|\bar{\mathcal{R}}([\tilde{\mathbf{P}}]^{t+1}) - \bar{\mathcal{R}}([\tilde{\mathbf{P}}]^t)\| \leq \epsilon_2$ or $t \geq T_2^{\max}$.

From (24), we know that the convex constraint $C1$ in problem (22) can be considered as a strengthening of the original nonconvex constraint $C1$ in (20). In other words, the feasible set defined in (22) is a *subset* of the true feasible set defined in (20). As a result, provided that the initial point $[\tilde{\mathbf{P}}]^0$ is feasible for the DCP (20), then all the iterates, $[\tilde{\mathbf{P}}]^t$ generated by iteratively solving the convex optimization problem (22) with the affine approximation in (21), always belong to the true feasible set defined by $C1$ and \mathcal{F} of (20). We summarize the proposed low-complexity solution as Algorithm 1, where we assume that an initial feasible point $[\tilde{\mathbf{P}}]^0$ of the DCP (20) is available (We will introduce how to obtain an initial feasible point in the following subsection.).

Remark that Algorithm 1 provides a low-complexity solution, in the sense that in each step a simple convex optimization problem is solved. The proposed Algorithm 1 converges to a local optimum after a few iterations, as can be observed from the simulation results.

C. Feasible Initial Point Searching Algorithm

Inspired by [40] (Sec. 11.4), [46], [47], we propose a feasible initial point searching algorithm, instead of an arbitrary point as in the conventional CCCP, to obtain the feasible $[\tilde{\mathbf{P}}]^0$ in Algorithm 1. The main advantage of the proposed new initialization method stems from the fact that, once the

proposed algorithm starts with a point in the feasible set of the DCP (20), all the iterates $[\tilde{\mathbf{P}}]^t$ generated by the algorithm remain within the original feasible set of the DCP (20). In addition, if the CCCP is initialized with a random (infeasible) point, the CCCP may fail at the first iteration due to the infeasibility of problem. However, the task of computing a feasible point of a nonconvex optimization problem, e.g., the problem (20), is NP-hard in general. This observation motivates the development of suboptimal, but low-complexity feasibility search procedures.

The proposed feasible initial point searching algorithm is based on similar iterative affine approximations of the originally nonconvex constraints as used in Algorithm 1, but with the following two modifications: a) the proposed searching algorithm starts with an arbitrary point $[\tilde{\mathbf{P}}]^0$; b) in the t -th iteration, instead of maximizing the spectrum efficiency of the networks as in problem (20), we maximize the slack parameter $s \in \mathbb{R}$, which can be regarded as an abstract measure of the constraint violations. The feasibility problem can then be expressed as the following convex program:

$$\begin{aligned}\max \quad & s \\ \text{s.t.} \quad & C1 : \bar{R}_n^b(\tilde{\mathbf{P}}) - \hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}}) \geq s, \forall n, \\ & \tilde{\mathbf{P}} \in \mathcal{F},\end{aligned}\quad (25)$$

where $\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}})$ is defined according to (23). If the current objective value s^{t+1} is zero, the algorithm stops; otherwise, the algorithm continues until convergence or until the maximum number of allowable iterations is reached. If no feasible point could be found with the proposed method, some admission control mechanisms can be adopted to reduce the number of MUs, which, however, is out of the scope of this paper. The proposed feasible initial point searching algorithm is summarized as Algorithm 2.

Note that a solution of problem (25) with $s = 0$ obtained is always feasible for the DCP (20). Conversely, if the proposed Algorithm 2 fails to provide a feasible point of problem (20), then this does not imply that this problem is infeasible since

Algorithm 2 The Proposed Feasible Initial Point Searching Algorithm

- 1: **Initialization:** Initialize maximum number of iterations T_3^{\max} and the maximum tolerance ϵ_3 ; Initialize the algorithm with a feasible point $[\tilde{\mathbf{P}}]^0$ and set the iteration number $t=0$.
 - 2: **Repeat:**
 - 3: Compute the affine approximation $\hat{R}_n^s([\tilde{\mathbf{P}}]^t, \tilde{\mathbf{P}})$ according to (21).
 - 4: Solve problem (25), and update $[\tilde{\mathbf{P}}]^{t+1}$ and $[s]^{t+1}$.
 - 5: Set $t = t + 1$.
 - 6: **Until:** $[s]^{t+1} = 0$ or $\|\mathcal{R}([\tilde{\mathbf{P}}]^{t+1}) - \eta([\tilde{\mathbf{P}}]^t)\| \leq \epsilon_3$ or $t \geq T_3^{\max}$.
-

Algorithm 2 operates only on a subset of the original feasible set of the DCP in (22).

The proposed Feasible Initial Point Searching Algorithm in Algorithm 2 together with the CCCP-based Algorithm 1 forms a two-step algorithm for solving the DCP in (20). In the first step, Algorithm 2 is applied to find a feasible point of the DCP in (20), instead of a random point. In the second step, the CCCP-based Algorithm 3 is applied, starting with the feasible point found in the first step. For convenience, the DCP solution algorithm in Algorithm 1 includes the feasible initial point searching algorithm in Algorithm 2 by default in the rest of this paper.

D. Overall Algorithm and Convergence Analysis

From Subsection IV-A and Subsection IV-B, it can be concluded that problem (10) can be solved by a two-tier iteration algorithm. In the first tier, the original problem is simplified by SCAM and variables transformation, and we can approach the original problem by updating the value of z_0 and choosing α and μ according to (11) in each iteration. In this way, we just need to solve one DCP in each iteration, which makes it possible to solve problem (10) easily. In the second tier, in order to solve the DCP, a CCCP-based iteration algorithm is proposed, in which the DCP is transformed to a convex problem by using Taylor expansion to approximate the non-convex constraint. In this way, the DCP can be solved by solving convex problems iteratively, which reduces the computational complexity significantly. The overall algorithm is summarized in Algorithm 3, where the feasible initial point is obtained by a similar algorithm with Algorithm 2.

Next, we analyze the convergence of our proposed power allocation algorithm in FD self-backhaul small cell networks with massive MIMO. The convergence proof is carried out in two parts: 1) The convergence proof of Algorithm 1 (the convergence behavior of Algorithm 2 can be inferred accordingly); 2) the convergence proof of Algorithm 3.

1) *Convergence Analysis of Algorithm 1:* We know that the point $[\tilde{\mathbf{P}}]^t$ is a feasible point of the convex optimization problem with concave objective function in (22), provided that the initial point $[\tilde{\mathbf{P}}]^0$ is feasible for the DCP in (20). As a consequence, the sequence $\{\mathcal{R}([\tilde{\mathbf{P}}]^t)\}$ monotonically increases as the iteration number t grows. Since the sequence $\{\mathcal{R}([\tilde{\mathbf{P}}]^t)\}$

Algorithm 3 The overall Algorithm

- 1: **Initialization:** Initialize maximum number of iterations T_1^{\max} and the maximum tolerance ϵ_1 ; Initialize the algorithm with a feasible initial $\tilde{\mathbf{P}}$, calculate the initial α and μ according to (12), and set the iteration number $t=0$.
 - 2: **Repeat:**
 - 3: Solve the problem (17) based on Algorithm 1 to obtain the current optimal $[\tilde{\mathbf{P}}]^{t+1}$.
 - 4: Update $\{[\alpha]^{t+1}, [\beta]^{t+1}\}$ according to (19) and (12).
 - 5: Set $t = t + 1$.
 - 6: **Until:** $\|\mathcal{R}([\tilde{\mathbf{P}}]^{t+1}) - \mathcal{R}([\tilde{\mathbf{P}}]^t)\| \leq \epsilon_1$ or $t \geq T_1^{\max}$.
-

is upper-bounded by transmission power limit (C2, C3 in problem (17)), the convergence of the sequence $\{\mathcal{R}([\tilde{\mathbf{P}}]^t)\}$, and thus the convergence of Algorithm 1 is guaranteed for any initial feasible point $[\tilde{\mathbf{P}}]^0$.

Moreover, since the objective function $\{\mathcal{R}(\tilde{\mathbf{P}})\}$ of problem (22) is strictly concave as we have proven in Subsection IV-B, the point $[\tilde{\mathbf{P}}]^{t+1}$, i.e., the solution of problem (22), is unique [40]. Hence, for any given initial feasible point $[\tilde{\mathbf{P}}]^0$, the entries of the two sequences, $\{\mathcal{R}([\tilde{\mathbf{P}}]^t)\}$ and $\{[\tilde{\mathbf{P}}]^t\}$, have a one-to-one correspondence. As a result, the monotone convergence of the sequence $\{\mathcal{R}([\tilde{\mathbf{P}}]^t)\}$ implies the convergence of the sequence $\{[\tilde{\mathbf{P}}]^t\}$, for any initial feasible point $\{[\tilde{\mathbf{P}}]^0\}$. Let $\tilde{\mathbf{P}}^*([\tilde{\mathbf{P}}]^0)$ denote the limit point of the sequence $\{[\tilde{\mathbf{P}}]^t\}$ with a feasible initialization $\{[\tilde{\mathbf{P}}]^0\}$ when the iteration number t goes to infinity, i.e., given the initial feasible point $\{[\tilde{\mathbf{P}}]^0\}$, we have

$$\tilde{\mathbf{P}}^*([\tilde{\mathbf{P}}]^0) \triangleq \lim_{t \rightarrow \infty} [\tilde{\mathbf{P}}]^t \quad (26)$$

In general, the limit point $\tilde{\mathbf{P}}^*([\tilde{\mathbf{P}}]^0)$ depends on the choice of the initial feasible point $\{[\tilde{\mathbf{P}}]^0\}$. For notational simplicity, we write the limit point as $\tilde{\mathbf{P}}^*$. Regarding the limit point $\tilde{\mathbf{P}}^*$, we have the following lemma.

Lemma 2: The limit point $\tilde{\mathbf{P}}^*$ of $\{[\tilde{\mathbf{P}}]^t\}$ is the solution of the following convex optimization problem:

$$\begin{aligned} \max_{\tilde{\mathbf{P}}} \quad & \mathcal{R}(\tilde{\mathbf{P}}) \\ \text{s.t.} \quad & \text{C1: } \bar{R}_n^b(\tilde{\mathbf{P}}) - \hat{R}_n^s(\tilde{\mathbf{P}}^*, \tilde{\mathbf{P}}) \geq 0, \forall n, \\ & \tilde{\mathbf{P}} \in \mathcal{F}, \end{aligned} \quad (27)$$

where the affine function $\hat{R}_n^s(\tilde{\mathbf{P}}^*, \tilde{\mathbf{P}})$ is obtained by replacing $\{[\tilde{\mathbf{P}}]^t\}$ with $\tilde{\mathbf{P}}^*$ in (24). Moreover, the limit point $\tilde{\mathbf{P}}^*$ satisfies all the constraints C1 in (27) with equalities, i.e.,

$$\bar{R}_n^b(\tilde{\mathbf{P}}^*) - \hat{R}_n^s(\tilde{\mathbf{P}}^*, \tilde{\mathbf{P}}^*) = \bar{R}_n^b(\tilde{\mathbf{P}}^*) - \bar{R}_n^s(\tilde{\mathbf{P}}^*) = 0. \quad (28)$$

Proof: By definition (26), the point $\tilde{\mathbf{P}}^*$ is the limit point of the sequence $[\tilde{\mathbf{P}}]^t$, hence the point $\tilde{\mathbf{P}}^*$ is a feasible point for the convex optimization problem (27) and no strictly better solution exists. What's more, as we have proven in IV-B, the objective function $\mathcal{R}(\tilde{\mathbf{P}})$ in (27) is strictly concave in the variable $\tilde{\mathbf{P}}$, so the solution of the problem (27) is unique [40] (Sec. 4.2), which means that the limit point $\tilde{\mathbf{P}}^*$ is the solution of problem (27). We prove the second part of the Lemma by contradiction. Assuming that the constraint C1 of the n -th SBS is not active, i.e., $\bar{R}_n^b(\tilde{\mathbf{P}}) - \hat{R}_n^s(\tilde{\mathbf{P}}^*, \tilde{\mathbf{P}}) > 0$, we can scale down the variable \tilde{P}_n^b to make the constraint active without

TABLE II: The simulation parameters

Simulation parameters	Value
The number of the MBS' antennas M	128
Path loss exponent	-3
Power spectral density of noise	-174dBm/Hz
Circuit power consumption P_C	160mW
Power amplifier efficiency $1/\rho$	38%
The power consuming weight w	10
The maximum transmission power of the MBS	46dBm
The maximum transmission power of SBSs	20dBm
The QoS requirement of MUs	2 bit/s/Hz

violating the other constraints, which makes it possible that the MUs or other SBSs' backhaul link can be allocated more power or lower interference level. Then the objective function will increase possibly, which contradicts the optimality of the point $\tilde{\mathbf{P}}^*$. Hence, it can be concluded that all constraints in (27) C1 are active at the point $\tilde{\mathbf{P}}^*$.

According to the Lemma 2, no matter how to choose the initial point $[\tilde{\mathbf{P}}]^0$, only if it is feasible, the final convergence point can be obtained by solving the problem (27). In other words, the limit point $\tilde{\mathbf{P}}^*$ is a stationary point of the DCP (17) [48]. Therefore, we have the conclusion that our proposed Algorithm 1 not only converges but also converges to a stationary point.

2) *Convergence Analysis of Algorithm 3:* We start the convergence analysis of Algorithm 3 with the following proposition.

Proposition 1: Algorithm 3 monotonically improves the value of the objective function at each iteration and converges. The solution obtained upon convergence satisfies the necessary optimality conditions.

Proof: Let $[\tilde{\mathbf{P}}]^t$, $[\alpha]^t$, and $[\mu]^t$ be the optimized values of the t -th iteration. Then, we have

$$\begin{aligned} \dots &\leq \bar{\mathcal{R}}([\tilde{\mathbf{P}}]^t) \stackrel{(a)}{=} \bar{\mathcal{R}}_{[\alpha]^t, [\mu]^t}([\tilde{\mathbf{P}}]^t) \stackrel{(b)}{\leq} \bar{\mathcal{R}}_{[\alpha]^t, [\mu]^t}([\tilde{\mathbf{P}}]^{t+1}) \\ &\stackrel{(c)}{\leq} \bar{\mathcal{R}}([\tilde{\mathbf{P}}]^{t+1}) \stackrel{(a)}{=} \bar{\mathcal{R}}_{[\alpha]^{t+1}, [\mu]^{t+1}}([\tilde{\mathbf{P}}]^{t+1}) \leq \dots, \end{aligned} \quad (29)$$

where the quality (a) is due to the fact that the relaxations of (13)(14)(15) are tight at the current SINR values according to (11) and (19), the inequality (b) is due to the fact that the maximization (17) is strictly concave, the inequality (c) follows from (18). For a finite set of transmit sum powers and channel gains, since the optimal spectrum efficiency is bounded above, the procedure must converge.

As a conclusion, the proposed power algorithm in FD small cell networks with massive MIMO converge well. We also evaluate its convergence performance by simulations in the next section.

V. SIMULATION RESULTS AND DISCUSSIONS

In this section, the effectiveness of our proposed FD self-backhaul scheme of small cell networks with massive MIMO will be demonstrated by Monte Carlo simulations, where the simulation results are averaged over 1000 droppings. In the simulations, we consider a $0.5\text{Km} \times 0.5\text{Km}$ square area covered by one MBS located in the center and some SBSs

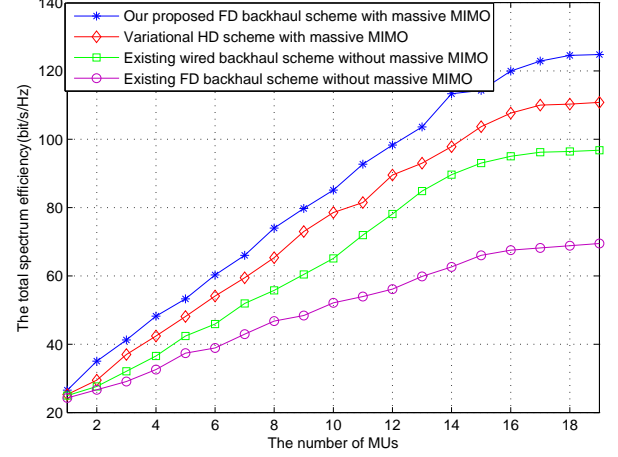


Fig. 3: The spectrum efficiency performance of different backhaul schemes with different numbers of MUs ($N = 4$).

(pico base stations are adopted) that are randomly deployed. The simulation parameters used are listed in Table II, which are similar to those in [49].

A. Performance Comparison with Existing Schemes

We evaluate the performance of the proposed backhaul scheme by comparing the following schemes: (a) a traditional wired backhaul scheme without massive MIMO [50], where the MBS schedules MUs in time domain; (b) an existing FD backhaul scheme without massive MIMO [11], where the MBS allocates the orthogonal frequency band to MUs and SBSs' backhaul link; (c) a variation of our proposed scheme with massive MIMO and half duplex (HD), where MBS can serve MUs and SBSs' backhaul link simultaneously, but each SBS receives and transmits data in different time slots. Each scheme has the similar system configurations as described above. In this subsection, we assume the self-interference is canceled incompletely ($\gamma = 10^{-5}$) [11].

1) *The effect of the number of MUs:* In Figs. 3, 4, 5, we compare the performance of different backhaul schemes with different numbers of MUs. As shown in Fig. 3, the spectrum efficiency performance of our proposed backhaul scheme outperforms the HD self-backhaul scheme with massive MIMO, the wired backhaul scheme without massive MIMO, and the FD self-backhaul scheme without massive MIMO. This is because our proposed backhaul scheme can take the advantage of both massive MIMO and FD backhaul. By massive MIMO technology, the backhaul DL of SBSs could be completed in the same frequency band at the same time with the MUs, rather than on the orthorhombic frequency band like in the schemes without massive MIMO, which can improve SE. By FD technology, the SBSs can receive backhaul data and transmit DL data simultaneously instead of in a time division pattern like in the HD backhaul scheme, which improves spectrum efficiency further. In addition, our proposed power allocation algorithm mitigates the inter-tier and intra-tier interference, which also has contribution to the

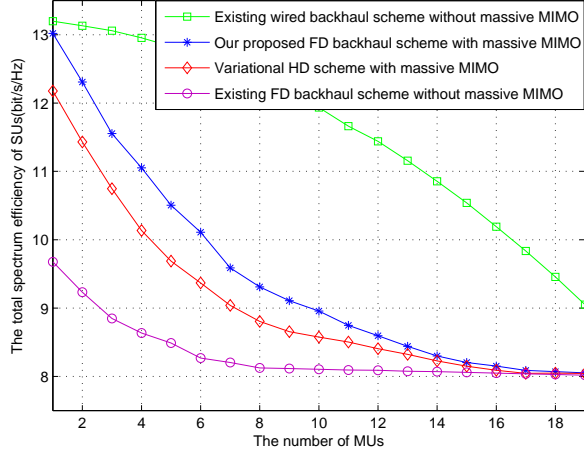


Fig. 4: The total spectrum efficiency of SUs of different backhaul schemes with different numbers of MUs ($N = 4$).

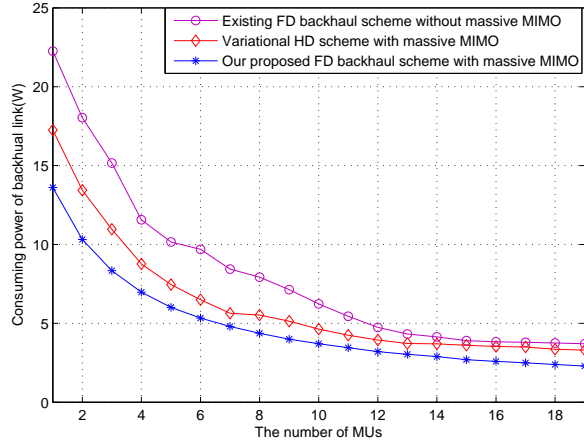


Fig. 5: The consuming power of backhaul link of different backhaul schemes with different numbers of MUs ($N = 4$).

high spectrum efficiency of our proposed backhaul scheme. What's more, it can be observed from Fig. 3 that the increasing rate of spectrum efficiency of all backhaul schemes becomes lower when the number of MUs reaches to some values. Actually, these inflection points mean the maximum number of serving MUs of different backhaul schemes. When the number of MUs exceeds the maximum number of serving MUs, the spectrum efficiency will continue to increase because of the multi-user diversity gain. Note that the maximum number of serving MUs is the highest in our proposed FD backhaul scheme with massive MIMO.

Fig. 4 and Fig. 5 show the change of the total spectrum efficiency of all SUs and the consuming power of backhaul link with the increase of the number of MUs, respectively. As shown in these two figures, for the wireless backhaul scheme, both the total spectrum efficiency of all SUs and the consuming power of backhaul link decrease when the number of MUs increases. The reason is that more MUs with better channel quality will encroach the power that should be

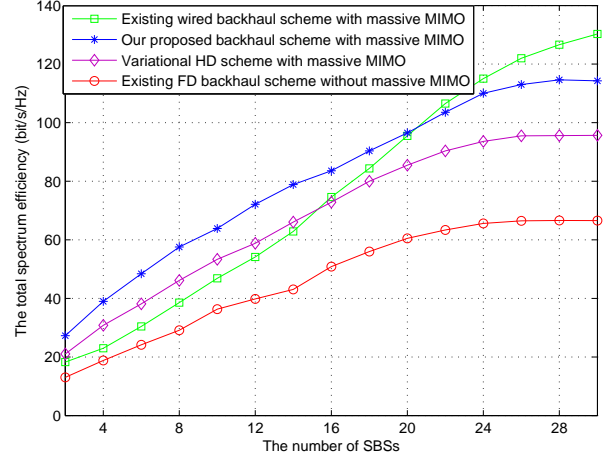


Fig. 6: The total spectrum efficiency of different backhaul schemes with different numbers of SBSs ($K = 4$).

allocated SBSs' backhaul link, and the spectrum efficiency of SUs will decrease. However, because of the limit of the lowest QoS requirement, the spectrum efficiency of SUs will not be zero. Comparing with FD backhaul scheme, the HD backhaul scheme only has half of time to transmit data to SUs and another half of time to receive data from MBS, so the spectrum efficiency of SUs of HD backhaul scheme will be low and the backhaul link will consume more power to reach the level of SE. In the non-massive MIMO backhaul scheme, besides power, the MBS needs to allocate orthogonal frequency band to SBSs' backhaul link, and then the available frequency band will reduce with the increase of the number of MUs, which results in the poor spectrum efficiency performance of SUs and consuming more backhaul power to satisfy the QoS. For the wired backhaul scheme, when the number of MUs increases, more MUs will appear near SBSs and they will aggravate the inter-tier interference to SUs, so the spectrum efficiency of SUs of wired backhaul scheme will decrease but the decreasing rate will be slow comparing with the wireless backhaul scheme. From Fig. 4 and Fig. 3, it can be observed that the spectrum efficiency of SUs of wired backhaul scheme will exceed other wireless backhaul scheme when the number of MUs is equal to about 6 but the total spectrum efficiency of the wireless backhaul scheme with massive MIMO is still higher than that of wired backhaul. This is because the MUs' spectrum efficiency gain obtained from massive MIMO is higher than the SUs' spectrum efficiency loss resulting from the less backhaul power when the number of MUs increases.

2) *The effect of the number of SBSs* : In Figs. 6, 7 and 8, we compare the total SE, spectrum efficiency of MUs and consuming power of backhaul link of different backhaul schemes with the increase of the number of SBSs, respectively. As shown in Fig. 6, the spectrum efficiency performance of all backhaul schemes firstly increases but the increasing rate decreases with the number of SBSs growing. The common reason is that the interference (inter-tier and intra-tier) level is relatively low when there are less SBSs and the interference will be serious when there are more SBSs. For wireless

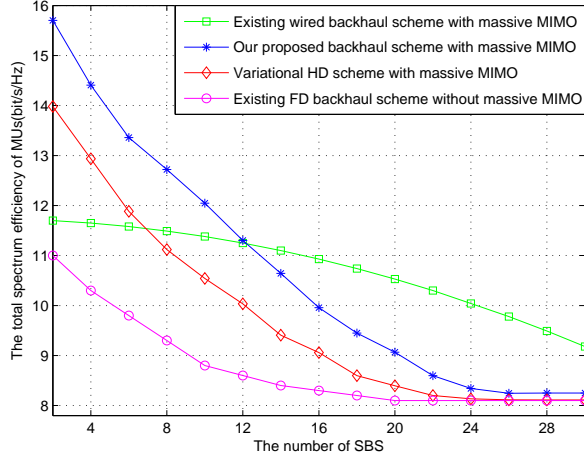


Fig. 7: The total spectrum efficiency of MUs of different backhaul schemes with different numbers of SBSs ($K = 4$).

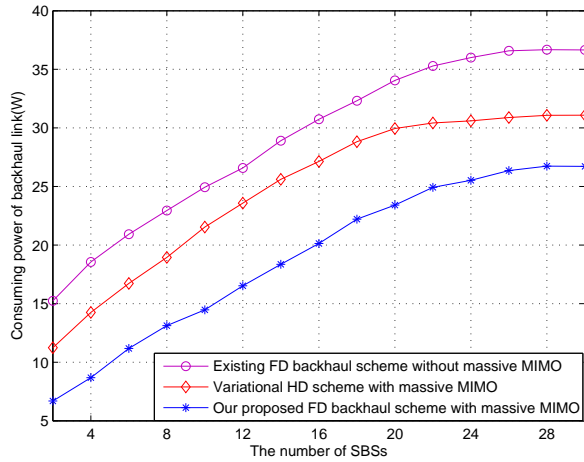


Fig. 8: The consuming power of backhaul link of different backhaul schemes with different numbers of SBSs ($K = 4$).

backhaul schemes, another main reason is that the backhaul power of each SBS will reduce with the increase of the number of SBSs due to the fact that the power of MBS is limited. This can be verified in Fig. 8, where consuming power of backhaul link will not grow any more when the number of SBSs is relatively large. At the same time, from Figs. 6, 7 and 8, we can find that the performance of our proposed self-backhaul scheme outperforms the HD self-backhaul scheme with massive MIMO and the FD self-backhaul scheme without massive MIMO. The reason is the same with that in Figs. 3, 4, 5. However, different from Fig. 3 and Fig. 4, the spectrum efficiency performance of our proposed FD self-backhaul scheme is better than that of the wired backhaul scheme without massive MIMO when the number of SBSs is less, but the spectrum efficiency performance of the wired backhaul scheme without massive MIMO will be better than our scheme when the number of SBSs increases to a relatively higher value. This is because the backhaul link needs to consume

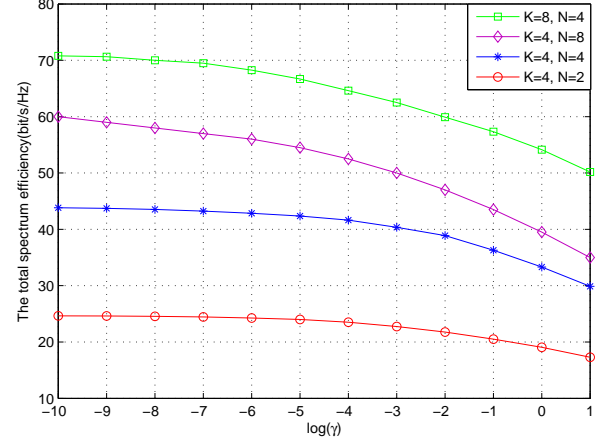


Fig. 9: The spectrum efficiency performance of different SI cancellation performance with different numbers of MUs and SBSs.

radio power in our scheme but it does not consume radio power in the wired backhaul scheme. When the number of SBSs is less, the backhaul consumes less power because of low interference level and lower QoS requirement of SUs. With the increase of the number of SBSs, the interference level and the QoS requirement of SUs will grow, which leads to consuming more backhaul power to satisfy the QoS requirement of SUs. This is also the reason why the spectrum efficiency of MUs keeps falling until the lowest QoS requirement is reached in Fig. 4. When the spectrum efficiency gain of SUs obtained by improving the number of SUs can not cover the spectrum efficiency cost of MUs because of the less available power, the spectrum efficiency performance of our proposed scheme will be worse than the wired backhaul scheme. Fortunately, our scheme is economic, and the wired backhaul is expensive. So a tradeoff exists between the network cost and network performance.

B. The Effect of Self-interference Cancellation Performance

In Fig. 9, we study the effect of self-interference cancellation performance on our proposed FD self-backhaul scheme. As shown in Fig. 9, the total spectrum efficiency will decrease with the increase of the value of γ and the decreasing rate becomes fast. This is because a larger value of γ means more serious self-interference, which results in the increase of consuming power of backhaul link. Consequently, the available power for MUs will be less, and the spectrum efficiency of MUs will decrease. At the same time, the MBS also will allocate less power to SBSs' backhaul link since the MBS will think the channel quality of those links is bad, which will influence the spectrum efficiency of SUs. As a result, the total spectrum efficiency performance will be bad when the self-interference is not cancelled perfectly. What's more, it can be observed from Fig. 9 that the total spectrum efficiency will be more sensitive to γ when there are more SBSs.

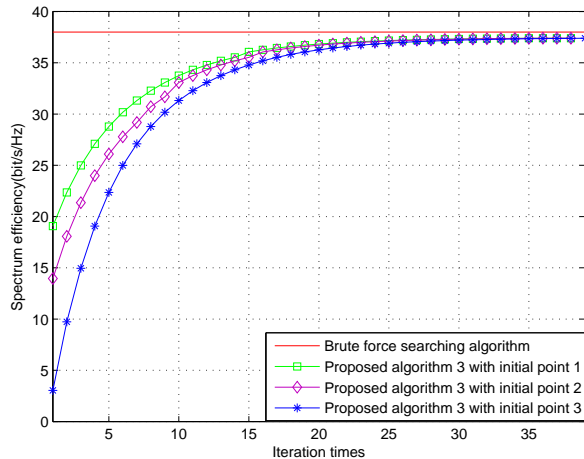


Fig. 10: The convergence of proposed overall algorithm with different initial point ($N = 4$, $K = 4$).

C. Convergence Illustration of the Proposed Algorithms

Fig. 10 shows the convergence behavior of the proposed power allocation algorithm in Algorithm 3 with different initial feasible points, as well as the optimal solution obtained by the brute force searching (BFS) algorithm, for the FD self-backhaul small cell networks including 4 MUs and 4 SBSs. As shown in this figure, we can observe the good convergence performance and robustness to initial points. No matter where the initial point is, our algorithm will converge. What's more, the gap between Algorithm 3 and the BFS algorithm is narrow after sufficient iterations, although our solution is not a global optimal solution, which means that the proposed Algorithm 3 is effective. Furthermore, it can be observed from this figure that a significant decrease of gap between Algorithm 3 and BFS algorithm can be found from the first iteration to the 10-th iteration. After the 10-th iteration, the gain of more iterations is still increasing but with less rate. Thus, a tradeoff exists between the acceptable utility value and iteration steps.

In Fig. 11, the convergence behavior of the proposed Algorithm 1 and two reference algorithms, i.e., the alternating optimization scheme of [51] and the EP-DCA of [41], are studied for the proposed FD self-backhaul small cell networks with 4 MUs and 4 SBSs. As shown in this figure, the proposed Algorithm 1 converges after approximately 20 iterations for any considered initial feasible point but the alternating optimization scheme of [51] and the EP-DCA of [41] converge after approximately 13 iterations, which implies that the convergence performance of the proposed Algorithm 1 is worse than that of the two reference algorithms. However, it is obvious that the optimization performance of our proposed Algorithm 1 outperforms the two reference algorithms, which is the reason why the performance of Algorithm 3 is close to the BFS algorithm.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed a self-backhaul scheme for small cell networks with massive MIMO and FD, which enables

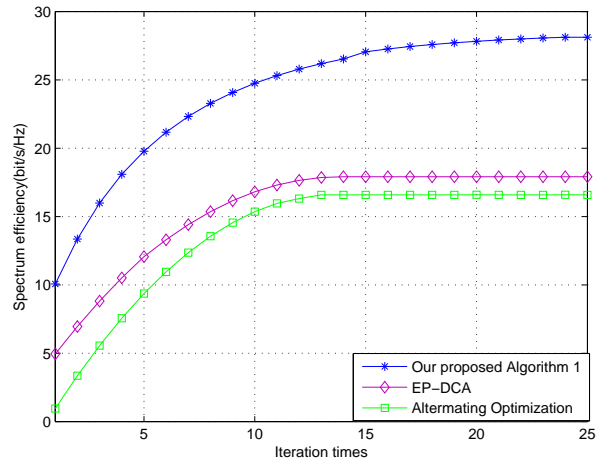


Fig. 11: The convergence of proposed CCCP-based algorithm ($N = 4$, $K = 4$).

the access of users and backhaul of SBSs simultaneously in the same frequency band. Furthermore, in order to mitigate the inter-tier and the intra-tier interference, we formulated the power allocation problem as an optimization problem, in which spectrum efficiency was taken as the optimization objective. Considering the high computation complexity for solving the non-convex optimization problem, we introduced the SCAM and an appropriate transformation of variables to transform equivalently the original problem into a DCP, which can be efficiently solved with local optimality using a CCCP-based algorithm. Simulation results showed that the proposed self-backhaul scheme by jointly using massive MIMO and FD technology is able to take the advantages of both massive MIMO and in-band FD communications. In addition, simulation results also demonstrated the effectiveness and good convergence performance of our proposed CCCP-based power allocation algorithm. Future work is in progress to consider the optimization of massive MIMO precoding matrix in our proposed scheme.

REFERENCES

- [1] Alcatel-Lucent, "The declining profitability trend of mobile data: What can be done? Assessing network costs and planning for sustainable revenue growth," Market Analysis, Tech. Rep., Dec. 2011.
- [2] S. Bu and F. R. Yu, "Green cognitive mobile networks with small cells for multimedia communications in the smart grid environment," *IEEE Trans. Veh. Tech.*, vol. 63, no. 5, pp. 2115–2126, June 2014.
- [3] S. Fortes, A. Aguilar-Garcia, R. Barco, F. Barba, J. Fernandez-luque, and A. Fernandez-Duran, "Management architecture for location-aware self-organizing LTE/LTE-a small cell networks," *IEEE Comm. Mag.*, vol. 53, no. 1, pp. 294–302, Jan. 2015.
- [4] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Commun. Surveys Tutorials*, vol. 17, no. 1, pp. 358–380, Firstquarter 2015.
- [5] J. Xu, J. Wang, Y. Zhu, Y. Yang, X. Zheng, S. Wang, L. Liu, K. Horneman, and Y. Teng, "Cooperative distributed optimization for the hyper-dense small cell deployment," *IEEE Comm. Mag.*, vol. 52, no. 5, pp. 61–67, May 2014.
- [6] C. Ranaweera, M. Resende, K. Reichmann, P. Iannone, P. Henry, B.-J. Kim, P. Magill, K. Oikonomou, R. Sinha, and S. Woodward, "Design and optimization of fiber optic small-cell backhaul based on an existing fiber-to-the-node residential access network," *IEEE Comm. Mag.*, vol. 51, no. 9, pp. 62–69, Sep. 2013.

- [7] S. Hur, T. Kim, D. Love, J. Krogmeier, T. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Wireless Comm.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.
- [8] R. Taori and A. Sridharan, "Point-to-multipoint in-band mmwave backhaul for 5G networks," *IEEE Comm. Mag.*, vol. 53, no. 1, pp. 195–201, Jan. 2015.
- [9] M. Mahloo, P. Monti, J. Chen, and L. Wosinska, "Cost modeling of backhaul for mobile networks," in *ICC Workshops*, June 2014, pp. 397–402.
- [10] D. Bladsjo, M. Hogan, and S. Ruffini, "Synchronization aspects in LTE small cells," *IEEE Comm. Mag.*, vol. 51, no. 9, pp. 70–77, Sep. 2013.
- [11] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Comm.*, vol. 32, no. 9, pp. 1637–1652, June 2014.
- [12] D. Hui and J. Axnas, "Joint routing and resource allocation for wireless self-backhaul in an indoor ultra-dense network," in *Proc. IEEE PIMRC*, Sep. 2013, pp. 3083–3088.
- [13] L. Erwu, J. Shan, S. Gang, and G. Luoning, "Fair scheduling in wireless multi-hop self-backhaul networks," in *Proc. AICT-ICIW*, Feb 2006, pp. 96–96.
- [14] G. Liu, F. Yu, H. Ji, V. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 2, pp. 500–524, Secondquarter 2015.
- [15] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Comm.*, vol. 13, no. 2, pp. 913–927, Feb. 2014.
- [16] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Comm.*, vol. 11, no. 12, pp. 4381–4393, Dec. 2012.
- [17] R. A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lahetkan-gas, and E. Tirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wireless Comm.*, vol. 22, no. 5, pp. 83–89, Oct. 2015.
- [18] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Areas Comm.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [19] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Comm.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [20] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell lab Tech. J.*, vol. 18, no. 2, pp. 5–21, Sep. 2013.
- [21] S. Lakshminarayana, M. Assaad, and M. Debbah, "Transmit power minimization in small cell networks under time average qos constraints," *IEEE J. Sel. Areas Comm.*, vol. 33, no. 10, pp. 2087–2103, Oct. 2015.
- [22] Y. Fei, V. W. S. Wong, and V. C. M. Leung, "Efficient QoS provisioning for adaptive multimedia in mobile communication networks by reinforcement learning," *Mob. Netw. Appl.*, vol. 11, no. 1, pp. 101–110, Feb. 2006.
- [23] C. Luo, F. R. Yu, H. Ji, and V. C. M. Leung, "Cross-layer design for TCP performance improvement in cognitive radio networks," *IEEE Trans. Veh. Tech.*, vol. 59, no. 5, pp. 2485–2495, 2010.
- [24] L. Ma, F. Yu, V. C. M. Leung, and T. Randhawa, "A new method to support UMTS/WLAN vertical handover using SCTP," *IEEE Wireless Commun.*, vol. 11, no. 4, pp. 44–51, Aug. 2004.
- [25] F. Yu and V. C. M. Leung, "Mobility-based predictive call admission control and bandwidth reservation in wireless cellular networks," in *Proc. IEEE INFOCOM'01*, Anchorage, AK, Apr. 2001.
- [26] Z. Li, F. R. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum sensing in cognitive radios," *IEEE Trans. Veh. Tech.*, vol. 59, no. 1, pp. 383–393, Jan. 2010.
- [27] F. Yu and V. Krishnamurthy, "Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff," *IEEE Trans. Mobile Computing*, vol. 6, no. 1, pp. 126–139, Jan. 2007.
- [28] R. Xie, F. R. Yu, H. Ji, and Y. Li, "Energy-efficient resource allocation for heterogeneous cognitive radio networks with femtocells," *IEEE Trans. Wireless Comm.*, vol. 11, no. 11, pp. 3910–3920, Nov. 2012.
- [29] S. Bu, F. R. Yu, Y. Cai, and P. Liu, "When the smart grid meets energy-efficient communications: Green wireless cellular networks powered by the smart grid," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3014–3024, Aug. 2012.
- [30] R. Xie, F. Yu, and H. Ji, "Dynamic resource allocation for heterogeneous services in cognitive radio networks with imperfect channel sensing," *IEEE Trans. Veh. Tech.*, vol. 62, no. 2, pp. 770–780, Feb. 2012.
- [31] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [32] Z. Jiang, A. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless Comm.*, vol. 14, no. 5, pp. 2868–2882, May 2015.
- [33] Z. Gao, L. Dai, W. Dai, and Z. Wang, "Block compressive channel estimation and feedback for FDD massive MIMO," in *Proc. IEEE INFOCOM WKSHPs 2015*, Apr. 2015, pp. 49–50.
- [34] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Comm.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
- [35] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Proc.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.
- [36] C. He, B. Sheng, P. Zhu, and X. You, "Energy efficiency and spectral efficiency tradeoff in downlink distributed antenna systems," *IEEE Wireless Comm. Let.*, vol. 1, no. 3, pp. 153–156, Jun. 2012.
- [37] J. Papandriopoulos and J. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.
- [38] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," in *Proc. IEEE INFOCOM*, vol. 2, June 2002, pp. 477–486.
- [39] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Comm.*, vol. 6, no. 7, pp. 2640–2651, July 2007.
- [40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2009.
- [41] L. T. H. An, "DC programming for solving a class of global optimization problems via reformulation by exact penalty," in *Proc. First Int'l Workshop on Global Constraint Optimization and Constraint Satisfaction*. Springer, 2003, pp. 87–101.
- [42] R. Horst and N. V. Thoai, "DC programming: overview," *J. of Optimization Theory and Application*, vol. 103, no. 1, pp. 1–43, 1999.
- [43] A. Smola, S. Vishwanathan, and T. Hoffman, "Kernel methods for missing variables," 2005.
- [44] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Advances in Neural Information Processing Systems*, 2009, pp. 1759–1767.
- [45] H. Li and T. Adali, "Complex-valued adaptive signal processing using nonlinear functions," *EURASIP J. on Advance in Signal Proc.*, vol. 2008, pp. 1–9, 2008.
- [46] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Processing*, vol. 60, no. 6, pp. 2962–2973, June 2012.
- [47] N. Bornhorst, M. Pesavento, and A. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Processing*, vol. 60, no. 1, pp. 221–232, Jan. 2012.
- [48] W. L. Winston and J. B. Goldberg, *Operations Research: Applications and Algorithms*. Duxbury press Boston, 2004, vol. 3.
- [49] L. Yan, B. Bai, and W. Chen, "On energy efficiency maximization in downlink MIMO systems exploiting multiuser diversity," *IEEE Comm. Let.*, vol. 18, no. 12, pp. 2161–2164, Dec. 2014.
- [50] Y. Yang, T. Quek, and L. Duan, "Backhaul-constrained small cell networks: Refunding and QoS provisioning," *IEEE Trans. Wireless Comm.*, vol. 13, no. 9, pp. 5148–5161, Sep. 2014.
- [51] Y. Jin and Y. Zhang, "Joint source and relay power optimization in multiuser cooperative wireless networks," in *Proc. 4th Int'l Symp. Comm., Control and Signal Processing (ISCCSP)*, Mar. 2010, pp. 1–4.